The Doctor Is In

Robert Borrelli

Until the late nineteenth century, mathematics was regarded as part of science and not as a separate discipline. Indeed, people such as Newton, Gauss, and Poincaré were comfortable in using high-level mathematics to explain phenomena in science. But eventually mathematics began to develop with little regard for applications, and departments of mathematics were founded at universities throughout the world. Mathematics returned to more practical concerns in the 1970s when traditional academic jobs in mathematics declined precipitously and math grads turned more and more toward jobs in industry.

Mathematicians in industry need skills different from those encountered in academia. Indeed, mathematicians in industry

- rarely choose their own problems to work on and are often not experts in problems they are assigned;
- almost always work in teams with engineers and scientists not of their own choosing, and must communicate across technical boundaries;
- are expected to take poorly defined problems and make sense out of them; and
- never have the luxury of unlimited time to work on a problem because projects in industry always have deadlines.

In 1973 the Mathematics Department at Harvey Mudd College (HMC) started an academic program to address the needs of industry-bound math grads. Its name, *Mathematics Clinic*, was borrowed from the HMC Engineering Department, which in the early 1960s created a program they called the *Engineering Clinic*. The name was chosen because they imagined it to be like the one in which medical interns get practice treating patients in a controlled and supervised environment. Although the word "clinic" has a remedial connotation, our *Clinic* program solicits real-world open-ended problems (called projects) from industrial concerns or government agencies that require hands-on teamwork by faculty and students over an academic year. Projects are tasks that require modeling, analysis, and validation and that offer significant educational value to students. Projects are assigned to a faculty member and a team of three or four students. Students do the work; the faculty member acts as an advisor. The "sponsor" is the industrial organization for whom we are consulting. A sponsor representative acts as liaison to the team. A team member acts as the project manager, who gives regular written and oral reports to the sponsor liaison. All *Clinic* project work takes place on campus, but sometimes the team travels to visit the sponsor.

In 1974 the Claremont Graduate University (CGU) established a *Mathematics Clinic* program, and these two Claremont Clinic programs cooperate in many ways for the benefit of all students in the Claremont Consortium. In the late 1970s the *Mathematics Clinics* were awarded a three-year NSF grant to train twelve postdocs to do *Clinic* projects; they all found jobs at year's end. Over two years in the mid-1980s, the Sloan Foundation supported eleven liberal arts math professors for a year-long *Math Clinic* experience. Currently, the norm is four projects per year at HMC and two projects per year at CGU. Since my knowledge concerns HMC *Clinic* operations, I will describe that program; CGU's program is similar.

Because "interns" in a *Math Clinic* project are undergrads, work is spread over an academic year—giving students time to learn the material required for their problem. Every *Clinic* project is a course, so students earn academic credit (but no pay), and the faculty advisor receives release time. HMC math students are required to do a capstone course of either a *Math Clinic* project or a senior

Robert Borrelli is emeritus professor of mathematics at Harvey Mudd College. His email address is `borelli@hmc.edu`.

thesis. Often students from other majors join *Math Clinic* teams. Because *Clinic* projects are such low-enrollment courses, the sponsor is required to pay a substantial fee (currently a standard $45K per project), and all rights to the project's results are assigned to the sponsor. The college receives 33% overhead per project to cover some college overhead costs and some faculty release time. Otherwise, the *Clinic* must cover its own expenses.

The *Clinic* is staffed part-time with a director, an administrative assistant, and a computer systems administrator. Operation of the program is aided by

- a detailed *Mathematics Clinic Handbook* for all participants;
- a carefully written and signed *Letter of Understanding*, which describes the project in enough detail to determine when it is completed (i.e., a "stopping rule");
- a signed *Confidentiality Agreement* to protect both the sponsor and the college;
- a *Clinic Advisory Committee* composed of representatives from industry, the *Clinic* directors, and the Director of Corporate Relations. This committee advises the directors on the operation of the *Clinic* program and, most importantly, aids the directors and the HMC Advancement Office in recruiting suitable projects.

Projects selected for the *Clinic* program must be educationally valuable to mathematics students; must not be of urgent value to the sponsor; must meet a reasonable prediction of success; and must be such that the sponsor can allocate a few hours per week for liaison time. Projects must also fit the time and resource constraints of an academic year. Prospective sponsors are asked to submit a list of three rank-ordered possible projects that they would support. The *Clinic* director and faculty make the final selection of projects using the criteria above. The *Clinic* always avoids half-year projects and is careful about accepting a software project that involves rewriting someone else's software code.

An organizational meeting at the beginning of the school year allows sponsors to present their projects so that students can see what mathematics is involved (statistics, operations research, computing, differential equations, etc.). Afterward, students express preferences before being assigned to particular projects. Each team reviews its work schedule often in order to finish the project by the academic year's end. *Clinic* teams must present weekly progress reports.

Clinic projects fall roughly into the categories listed below (along with examples):

- Algorithm Development
  - **Fair, Isaac** (Rule-based neural nets)
  - **Aerojet** (Multiple hypothesis testing)
- Optimal Design
  - **Teledyne** (Design of a hybrid microprocessor)
  - **AlliedSignal** (Design of a turbocharger blade)
- Modeling/Optimization
  - **B2 Division Northrop** (Maximization of EM energy absorption)
  - **Chevron** (Prediction of abnormally high underground water pressures)
- Smart Software
  - **Jet Propulsion Laboratory** (Formal verification methods based on math logic)
  - **Texas Instruments** (Synthetic Speech Listener)
- Software Design
  - **Bank of America** (Design of contact-sensitive GUI components)
  - **CODEE Consortium** (Design of a Java version of an ODE solver)

The *Clinic* director passes along this advice to faculty advisors:

- be professional in all relations with the sponsor;
- don't expect the sponsor to always follow your advice;
- be aware of the realities of student commitments;
- be careful to adhere to the Confidentiality Policy; and
- expect to have personnel problems (it happens in industry, too).

Clinic projects benefit sponsors in many ways. The *Clinic* can be regarded as an extension of in-house R & D at a much more reasonable cost. *Clinic* work provides the sponsor's staff contact with academic applied mathematicians (resulting sometimes in consulting opportunities). Teams have access to all library facilities and all computer facilities at HMC. And last but not least, the *Clinic* provides sponsors with a chance to see how possible employees might perform on the job.

The *Clinic* also benefits students in many ways. For example, students learn valuable techniques they may not have seen in a course and learn how to work together in teams to accomplish a common objective. Students learn how to work under time and resource constraints and how to set realistic schedules. They also learn how to write clear and concise technical reports and how to give professional-level talks and briefings. By planning day-to-day activities for the team, project managers gain valuable management experience.

Students often receive job offers from the sponsor at the project's conclusion.

The academic year ends with all the forty or so Mathematics and Engineering Clinics projects presenting their results at a professional-level meeting called Projects Day. Projects Day attracts some 400 people and presents a great opportunity to recruit projects for the following year.

To create a viable *Clinic* program, these questions must be answered:

- What makes a good *Clinic* project?
- How well can undergrads handle open-ended projects?
- What exactly is the role of the faculty advisor?
- How are faculty advisors recruited?
- How are sponsors recruited?

Several indicators show that HMC has successfully answered these questions: To date (1973 –2010), the HMC *Mathematics Clinic* has had 130 projects from 58 sponsors involving 553 students. A large number of sponsors return year after year, and the Math Clinic has inspired other institutions to create similar programs, among them San Jose State University, the University of South Australia, and the Institute for Pure and Applied Mathematics (IPAM) at UCLA.

For detailed project abstracts and other information about the HMC *Math Clinic* program, visit `http://www.math.hmc.edu/clinic`. For further information contact the HMC *Math Clinic* director, Susan Martonosi, at `martonosi@math.hmc.edu` (or at CGU contact Ellis Cumberbatch at `ellis.cumberbatch@cgu.edu`). Institutions wanting to start a *Clinic* program are invited to contact either *Math Clinic* director for advice. Of course the best introduction is for an interested faculty member to go through a complete *Clinic* cycle while on sabbatical leave for a year!

Book Review

The Cult of Statistical Significance

Reviewed by Olle Häggström

The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives
S.T. Ziliak and D. McCloskey
University of Michigan Press, 2008
US$26.95, 352 pages
ISBN-13: 978-0472050079

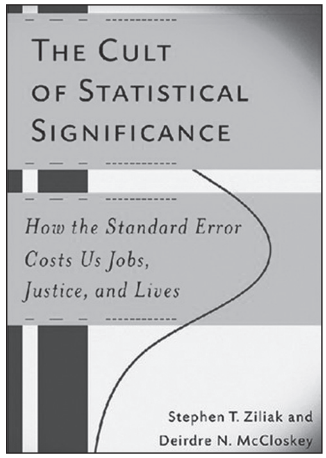
There are excellent evolutionary reasons why we humans have far-reaching abilities to observe the world around us and to draw sensible conclusions about it. But evolution is very far from a perfect optimization algorithm, so it should come as no surprise that our cognitive capacities, too, are far from perfect. One example is our strong tendency to deduce patterns from meaningless noise. Another is our inclination toward overconfidence in our conclusions, as evidenced by studies showing how in certain kinds of situations we are wrong about 40 percent of the time about conclusions we claim to be 98 percent sure about [AR].

The scientific method can be seen as an organized attempt to overcome such "bugs" in our search for accurate knowledge about the world around us. One ingredient, which during the course of the twentieth century has permeated all of science to the extent that it is nowadays recognized as indispensable, is mathematical statistics, which helps researchers distinguish between pattern and noise and to quantify how much confidence in our conclusions the data warrant.

There can hardly be any doubt that this development has been of immense benefit to science. All the more interesting, then, that two prominent economists, Stephen Ziliak and Deirdre McCloskey, claim in their recent book *The Cult of Statistical Significance* [ZM] that the reliance on statistical methods has gone too far and turned into a ritual and an obstacle to scientific progress.

A typical situation is the following. A scientist formulates a *null hypothesis*. By means of a *significance test*, she tries to falsify it. The analysis leads

Olle Häggström is professor of mathematical statistics at Chalmers University of Technology in Gothenburg, Sweden. His email address is `olleh@chalmers.se`.



to a *p-value*, which indicates how likely it would have been, if the null hypothesis were true, to obtain data at least as extreme as those she actually got. If the *p-value* is below a certain prespecified threshold (typically 0.01 or 0.05), the result is deemed *statistically significant*, which, although far from constituting a definite disproof of the null hypothesis, counts as evidence against it.

Imagine now that a new drug for reducing blood pressure is being tested and that the fact of the matter is that the drug does have a positive effect (as compared with a placebo) but that the effect is so small that it is of no practical relevance to the patient's health or well-being. If the study involves sufficiently many patients, the effect will nevertheless with high probability be detected, and the study will yield statistical significance. The lesson to learn from this is that in a medical study, statistical significance is not enough—the detected effect also needs to be large enough to be *medically significant*. Likewise, empirical studies in economics (or psychology, geology, etc.) need to consider not only statistical significance but also economic (psychological, geological, etc.) significance.

A major point in *The Cult of Statistical Significance* is the observation that many researchers are so obsessed with statistical significance that they neglect to ask themselves whether the detected discrepancies are large enough to be of any subject-matter significance. Ziliak and McCloskey call this neglect *sizeless science*. They exemplify and discuss instances of sizeless science in, among other disciplines, medicine and psychology, but for obvious reasons they focus most of their attention on economics. In one study, they have gone over all of the 369 papers published in the prestigious journal *American Economic Review* during the 1980s and 1990s that involve regression analysis. In the 1980s, 70 percent of the studied papers committed sizeless science, and in the 1990s this alarming figure had increased to a stunning 79 percent. A number of other kinds of misuse of statistics are considered in the same study, with mostly equally depressing results.

One particular error, which every teacher of mathematical statistics is painfully familiar with, is to conflate the probability of the observed data given the null hypothesis with the probability of the null hypothesis given the data (the latter cannot, of course, be obtained unless we resort to Bayesian statistics, a framework that is still rare in the fields under study). This error, known as the fallacy of the transposed conditional, is discussed in the book but does not appear as a separate item in the *American Economic Review* literature study.

The Cult of Statistical Significance is written in an entertaining and polemical style. Sometimes the authors push their position a bit far, such as when they ask themselves: "If null-hypothesis significance testing is as idiotic as we and its other critics have so long believed, how on earth has it survived?" (p. 240). Granted, the single-minded focus on statistical significance that they label sizeless science is bad practice. Still, to throw out the use of significance tests would be a mistake, considering how often it is a crucial tool for concluding with confidence that what we see really is a pattern, as opposed to just noise. For a data set to provide reasonable evidence of an important deviation from the null hypothesis, we typically need *both* statistical *and* subject-matter significance.

The book also offers a short history of the significance test. Here Ziliak and McCloskey take their polemical style to even further heights in their portrayal of William Gossett (inventor of Student's t-test, the most widely used of all significance tests) as a hero and an angel and of Ronald Fisher (the father of modern mathematical statistics, who arguably did more than anyone else to give significance testing the central role it has today) as pretty much the devil himself. For instance, they make no attempt at concealing their schadenfreude when quoting what Robert Oppenheimer (allegedly) had said upon Fisher's arrival in Berkeley in 1936: "I took one look at him and decided I did not want to meet him" (p. 222).

To sum up, if statistical practice in the empirical sciences is as bad as the authors say, what should be done? No easy fix is offered, but they do advocate a larger degree of pluralism among statistical methods. Here, one would have liked to see them address the danger that this might lead to an increase in a particular kind of misuse of statistics: to tune the choice of statistical approach to the particular data that were obtained. Many of the authors' comments seem to imply a commitment to the Bayesian paradigm, but it is not clear whether they are really aware of this. In any case they never explicitly step out of the Bayesian closet.

References

[AR] M. Alpert and H. Raiffa, A progress report on the training of probability assessors, *Judgment under Uncertainty: Heuristics and Biases* (D. Kahneman, P. Slovic, and A. Tversky, eds.), Cambridge University Press, 1982.

[ZM] S. T. Ziliak and D. McCloskey, *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*, The University of Michigan Press, Ann Arbor, MI, 2008.